Wireless Pers Commun (2018) 100:283–309
https://doi.org/10.1007/s11277-017-5066-6

Progressive Streaming of Video Data for Traffic Surveillance

Shivprasad P. Patil[1] · Rajarshi Sanyal[2] · Ramjee Prasad[1]

Published online: 22 November 2017

Abstract Traffic surveillance has been one of the essential attributes in smart city concept. Nowadays, in such applications rotating camera is preferred over static camera. Motivation behind this substitution is to reduce the cost of data transmission and Total of cost of ownership. To design an optimal and performant wireless 'smart city area network' for video surveillance systems, this paper focuses on some key areas, namely, transmission efficiency, lossless video data coding, data congestion, edge computing at transmission nodes. The end objective is to achieve high quality received video stream in spite of compressed data transmission. Some research initiatives in this domain are pertinent. For example, Structural Similarity Index (SSIM) based rate distortion optimization is an effective tool in enhancing the perceptual video quality in wireless environments. However, prevailing system does not consider the network congestion conditions, affecting quality of received video. Also, effect of distortion introduced by 'channel noise' is unattended. This motivated us to propose a new dual metric traffic control mechanism, wherein both metrics i.e. distortion and data congestion are considered. It is based on an 'improvised SSIM' method which incorporates the 'Rate of allocation' algorithm as a function. Experimental results unveil that the proposed traffic control using similarity index under noise diversity can achieve better video quality and more data throughput.

Keywords Automated traffic surveillance · Video streaming · Error resilience · Rate of allocation · Noise distortion · Structural similarity index

✉ Shivprasad P. Patil
spptil1212@gmail.com

Rajarshi Sanyal
rajarshi.sanyal@bics.com

Ramjee Prasad
ramjee@btech.au.dk

1 Department of Business Development and Technology, Aarhus University, Herning, Denmark

2 Belgacom International Carrier Services, Brussels, Belgium

1 Introduction

In various real time applications, traffic surveillance has gained its interest due to rapid growth in vehicular traffic density and the criticality in monitoring it. It has been extremely important to monitor the traffic flow in real time. In automation of traffic monitoring, the captured traffic video data need to be monitored at a remote monitoring unit. In legacy implementations, multiple static video cameras were installed at traffic junctions or on highways to capture the traffic data of moving vehicles. Video streams from each of these cameras were fed to the centralized monitoring network. Therefore, it is needed to route the data to the monitoring station at a faster rate and with maximum visual accuracy. In the transmission process, video frames are forwarded in a multihop manner, where each link point located at a distance route the data to the next destination. In this routing process, probability of congestion arises due to continuous and volumetric data streaming. Hence, such nodes need to be made 'congestion control' to have higher throughput.

More recently, there is a trend to replace the static cameras with rotating video camera with an endeavor to reduce the data transmission cost and reduce the hardware requirements. In earlier work [1], we were able to achieve the suppression of redundant data along with rejection of false positives of motion element by using a recurrent block matching approach. This reduces transmission data and thereby transmission cost. As the video cameras are installed at remote locations demanding a wide observation area, it is required to consolidate all observations at a centralized location for further analysis. The information from individual cameras needs to be conveyed over a wireless network to the central monitoring office. The challenges for setting up such a wireless network are bountiful. Some of the important are video data throughput, nonlinear nature of channel and received data accuracy. To reduce the transmission overheads, this paper intend to compress the video data at the node itself. This is done by detecting a moving object with the intention to suppress repetitive background information. To achieve the objective of error resilient coding in video streaming, a rate distortion optimization (RDO) using the structural similarity index metric (SSIM) was proposed in [2]. This solution optimizes the wireless video streaming operation. It also defines a Lagrange optimization method for video coding process. The said approach improves the conventional model of SSE-based error-resilient RDO for wireless video streaming, wherein less computing resources are needed. Though SSIM-RDO is more accurate (w.r.t. transmitted and received values of pixel data) and uses lesser computing resources [2], yet it lacks functionalities like error control. Moreover, no emphasis on the transmission model and network topology was laid. This is essential to know to avoid the errors due to congestion. If a hop based network is considered for communication, the probability of developing congestion proliferates as each hop can contribute to its part of the queue build up. This is depicted in Fig. 1.

Congestion may be caused by many factors. If, packets begin arriving on multiple ingress channels and all are multiplexed towards a common egress channel, then the aggregator node can experience higher load compared to the other nodes in the hop. Plus each node adds on to its own traffic. So it is imperative that if there is insufficient buffer, packets will be lost. Therefore congestion control needs to be addressed in such a network. In [3], an adaptive queue mechanism based on context modelling is proposed, wherein, context-aware adaptive active queue management (CA-AQM) approach of controlling the flow of packet based on network condition and video compression characteristics is proposed. In this queue management process, the flow of a video sequence is controlled by dropping or forwarding the data based on the queue size. Further it is observed, the distortion at the channel level are highly variant in nature, which introduces dynamic noise.

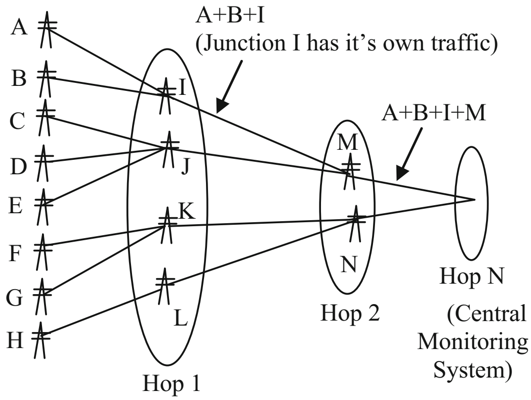


Fig. 1 Data traffic flow from source to destination

The work carried out in [4] defines an approach towards provision of rate distortion optimization using SSIM metric under channel variant conditions.

However, all the conventional methods developed were confined to node level. In the network level, the blockage factor at link node could reduce the traffic flow, which brings the rate allocation efficiency down. Hence, the conventional model of SSIM-RDO system needs to be optimized with dual optimization metric of distortion variability and rate allocation considering similarity index as observing metric. In this work, a novel approach is presented for data rate allocation with dual metrics, distortion and network congestion for wireless network with bit error rate (BER) .001. Wherein, non variable and variable channel conditions are considered. The rest of the paper is organized as follows. Section 2 outlines the work done so far on the traffic surveillance in a wireless environment and its real time applications. Section 3 presents the 'State of the Art' of the SSIM-RDO approach for error-resilience coding and queue management technique. In Sect. 4, the proposed approach with data flow control with consideration of distortion and data traffic congestion, under non-variable and variable channel conditions, is explained. The experimental results for the developed approach are presented in Sect. 5. Section 6 presents the conclusions and future scope for the present work.

2 Past Work

There has been a considerable volume of research in the domain of video transmission for surveillance systems. This section lists out some of these researches that focus on the algorithmic approaches and the methods that are available currently. In [5], an analysis of power consumption in video coding based on the constant bit rate over the 3G service is presented. The approach of radio resource utilization based on the control transition state machine was proposed for a 3G network. A relay based communication for video transmission was presented in [6], wherein receiver-based solution with video transmission decoupled from relay node selection (REDEC) was developed. The approach uses receiver modeling considering the excessive collision and overhead due in the exchange of video packets at a high frequency. In [7], a source rate control technique for video streaming is controlled over a wireless channel by restoring on a reduced reference (RR) quality estimation approach. The process extracts the content feature of the video sequence, which is transmitted with the video sequence. The transmitted features are used to analyze the quality of the video received. The source rate is then controlled to achieve the objective of a higher throughput and visual quality. But, RR approach of quality estimation is less accurate and efficient than full reference (FR), like SSIM. In [8], an integrated model of multi source video monitoring for traffic surveillance is proposed. A block partitioning of the video data into macroblocks for efficient streaming is proposed. The approach analyzes the impact of different multipath conditions for data streaming over vehicular ad hoc networks (VANET), taking various network parameters into consideration. But, no attention on network throughput is given in this work. To add intelligence to the video streaming operation in [9], a neuro-fuzzy modeling for MPEG-4 video transmission over IEEE 802.15.4.zigbee wireless device was proposed. The approach defines two schemes for monitoring the input and output of the data storage in traffic regulation. The approach controls the bit rate coding for traffic application to overcome the picture loss quality in MPEG-4 video coding over the zigbee applications only. In [10], a real time vision system, for automatic traffic monitoring was presented. The system was designed for automated capturing and processing of images from the pre-calibrated cameras. The method was developed under the framework of TRAVIS (Traffic Visual monitoring) project work. Here, more focus is given on video processing aspects than data communication issues. Towards the seamless transmission in [11], a new protocol based on network-based localized mobility management group working of the internet engineering task force was proposed. The protocol minimizes the issues for network switching under the mobility scenario. The protocol focuses on mobility management to minimize latency, jitter, and packet loss in the video streaming application. In this work, more attention is given on avoidance of network delay than visual experience at receiving end. A cross layer approach of video streaming over a wireless network is presented in [12]. The scheme proposes a rate adaptation for data link and a physical layer, whereas the quality adaptation in the application layer. The rate adaptation is used for the adjustment of the allocated rate, where as the quality adaptation scheme controls the video quality offered. In this work, effect of congestion due to network conditions is not considered for data rate adaptation. In [13], to improve the video quality a joint selection of the quantization offset is presented. The statistical distribution of the transformed coefficient in an encoded video sequence is developed using a Laplacian model. The multi level optimization solution results in an optimal path selection for lower path failure probability. The seamless transmission of a video sequence using H.264/AVC was developed. Here, distortion metric to check the video quality at receiving end is not matching with requirements of the HVS. In [14], the effective channel bandwidth and the current channel state are analyzed for the automatic repeat request error controlling operation. The constraint of buffer and end-to-end delay is considered as a trans-coding parameter. It is illustrated that the approach of trans-coding results in improved video picture quality. But, in this work, rate adaptation is done for H.263 system and visual quality metrics used is PSNR. A scalable mode of video coding is presented in [15]. The approach of vehicle monitoring, in the multi-Hop communication model is presented. However, perceived video quality of experience is compromised. In [16] a motion based compensation of rate allocation using SSIM metric was outlined. The method proposes a data rate compensation for transmission. However, the source side encoding is not been evaluated for rate allocation. In [17] a complex wavelet based coding for SSIM based rate allocation is suggested. However, effect of dynamic channel conditions is not considered in the work.

3 State of the Art (SSIM Measure for RDO)

The use of SSIM-RDO for error resilience coding in video streaming is very popular. The reason behind this, SSIM outperforms traditional methods, such as peak signal to noise ratio (PSNR) and mean squared error (MSE), which have proven to be consistent with human perception. As proposed algorithm of this paper uses the concept of SSIM-RDO, it is imperative to know the essence of the same.

3.1 SSIM-Based RDO Formulation Based on SSE-RDO

In [2], the SSIM-RDO based error-resilient scheme for H.264/AVC is presented. To improve the wireless video streaming performance, a numerical relation was derived through the Lagrange method to obtain minimum distortion, wherein the SSIM is used as a distortion metric. A low-complexity Lagrange multiplier for SSIM-based RDO for error-free coding is derived initially. The SSIM-based decoding for distortion minimization is introduced in the encoder to formulate error resilient video coding. For the distortion optimization in video streaming, similarity index measurement is a qualitative measuring index parameter for video streaming.

In the SSIM based video coding, the encoding process can be determined by reaching the best trade-off between the coding bits amount and the obtained video quality. This problem can be modeled as;

$$\min_{\{m\}}\{D\} \Rightarrow R_e \le R_l \tag{1}$$

Which indicates that the video encoder should minimize the apparent distortion ‘D’ with the number of encoding bits amount ‘Re’, following the constraint of bits amount ‘Rl’ by selecting the appropriate encoding mode ‘m’ [2]. In video streaming, the Lagrange optimization approach is used to make the objective as;

$$\min_{\{m\}}\{J\} = D + \lambda . R_e \tag{2}$$

where ‘J’ is Lagrange cost and ‘λ’ is the Lagrange multiplier for RDO. Normally in Lagrange optimization, the distortion metrics such as SSE and SAD are used as measures of video quality. However, these methods are not able to model the quality that is perceived by the Human Visual System (HVS). Human perception is naturally adapted to luminance, contrast and structure in an image, which is basis for SSIM.

The SSIM metric is measured as a similarity metric, for the original (I) and the distorted video data (I'), which is used with the Lagrange multiplier to control the encoding bits. The SSIM metric is given as a correlative factor defined as a function of mean, standard deviation, cross correlation for the two video data.

The SSIM metric is defined by;

$$\mathrm{SSIM}(I, I') = \frac{(2\varphi_I \varphi_{I'} + A)\left(2\tau_{I,I'} + B\right)}{I^2 + I'^{(2+A)}\left(\tau_I^2 + \tau_I'^2 + B\right)} \tag{3}$$

where ‘φ_I’, ‘τ_I’ and ‘$\tau_{I,I'}$’ are the mean, standard deviation, and cross correlation between the two video data respectively. ‘A’ and ‘B’ are used as a stabilizing parameter for the means and variances to set near to zero.

To optimize the distortion metric, the Lagrange optimization is defined by,

$$\min_{\{m\}}\{J\} = D_{SSIM} + \lambda_{SSIM} \cdot R_e \tag{4}$$

where 'DSSIM' denotes the SSIM-based distortion and 'λ_{SSIM}' is the Lagrange multiplier for the SSIM-based RDO. To optimize the rate allocation the Lagrange operator is to be suitably chosen to fix an optimal rate.

It is worth to note; in SSIM based method the Lagrange multiplier is theoretically derived from sum of squared (SSE) based Lagrange optimization process, thereby reducing mathematical complexity.

$$\lambda_{SSIM} = \frac{-D_{SSIM}}{R_e} = -\frac{\left(\frac{D_{SSE}}{f}\right)}{R_e} = -\frac{1}{f} \cdot \frac{D_{SSE}}{R_e} = \frac{\lambda_{SSE}}{f} \tag{5}$$

Thus, for the SSIM-based Lagrange optimization process, it can be modeled by only scaling the existent SSE-based Lagrange optimization formulation with a fixed factor 'f' as [2];

$$\min_{\{m\}}\left\{\frac{\lambda_{SSE}}{f}\right\} \text{with } \frac{\lambda_{SSE}}{f} = \frac{D_{SSE}}{f} + \frac{\lambda_{SSE}}{f} \cdot R_e \tag{6}$$

3.2 Approach A: SSIM-Based Error-Resilient Video Coding Under Non Variant Channel Condition

For the error distortion minimization in this approach, to provide the network optimality, the video coding layer (VCL) and the network abstraction layer (NAL) are designed for the H.264/AVC video coding standard. The VCL operates for the compression process whereas; the NAL operates at the network level to provide proper allocation of resources. For wireless communication, the transmission channel is time-varying and erroneous in nature. For the minimization of an error during signal propagation, an independent channel model is used. Knowing the bit error rate (BER) of the transmission channel, the packet loss probability 'ρ' for a NAL unit containing 'L' bits is related as;

$$\rho = 1 - (1 - ber)^L \tag{7}$$

During the encoding process, the video streams are divided into frame slices represented as $s_{n,m}$. For the mth slice in the nth frame, the BER is defined by $ber_{n,m}$(which is the channel BER for the transmission of the mth slice of the nth frame), and $\rho_{n,m}$ is the packet loss rate for slice $s_{n,m}$. The Lagrange multiplier 'λ_{SSIM}' was adjusted to achieve the objective of the error resilient to a minimum. The Lagrange multiplier was developed based on the distortion metric 'D_{SSIM}'. Since the distortion estimation is conducted at the encoder end, a module in the encoder is added, to simulate the decoding process with the help of the acknowledgement message, which informs the encoder whether the transmitted packet is received by the receiver or not. For an acknowledgement message of 'nr' frame received by the encoder while encoding the 'n' frame, the encoding information is stored and the added decoding unit decodes the 'nr' frame and gets the expected decoded frames from the nr + 1 to the n − 1 frame. Given the decoded reference frames or the expectations of the decoded reference frames, the pixel values were obtained. Further, the expected decoding distortion is estimated by;

$$E\{DSSIM_{n,m,k}\} = 1 - \rho_{n,m} \cdot SSIM\left(b_{n,m,k}, b^{e-c}_{n,m,k}\right) - \left(1 - \rho_{n,m}\right) \cdot SSIM\left(b_{n,m,k}, b^{n-l}_{n,m,k}\right) \tag{8}$$

where $b_{n,m,k}$, $b^{e-c}_{n,m,k}$ and $b^{n-l}_{n,m,k}$ indicate the original MB, the error concealed MB with packet loss and the decoded MB without packet loss, respectively. The proper adjustment of the Lagrange multiplier based on the distortion metric is evaluated as;

$$\begin{aligned}
\lambda'_{SSIM} &= \frac{DSSIM(R_e)}{R_e} \\
&= \frac{\left(1 - \rho_{n,m} \cdot SSIM\left(b_{n,m,k}, b^{e_c}_{n,m,k}\right) - \left(1 - \rho_{n,m}\right) \cdot SSIM\left(b_{n,m,k}, b^{n_l}_{n,m,k}\right)\right)}{R_{e_{n,m,k}}} \\
&= \frac{\left(\rho_{n,m} \cdot SIM\left(b_{n,m,k}, b^{e_c}_{n,m,k}\right)\right)}{R_{e_{n,m \cdot k}}} + \frac{\left(\left(1 - \rho_{n,m}\right) \cdot SSIM\left(b_{n,m,k}, b^{n-l}_{n,m,k}\right)\right)}{R_{e_{n,m \cdot k}}}
\end{aligned} \tag{9}$$

Approximately, represented as;

$$\frac{\left(\rho_{n,m} \cdot SIM\left(b_{n,m,k}, b^{e-c}_{n,m,k}\right)\right)}{\partial R_{n,m.k}} = -\frac{\left(\left(1 - \rho_{n,m}\right) \cdot SSIM\left(b_{n,m,k}, b^{n_l}_{n,m,k}\right)\right)}{R_{n,m.k}} \approx \lambda_{SSIM} \tag{10}$$

where 'λ_{SSIM}' indicates the Lagrange multiplier for the RDO in the error-free environment. When the Lagrange multiplier, 'λ'_{SSIM}' is adjusted to be smaller than 'λ_{SSIM}' the error-resilient RDO will select more intra-coded bits to restrain the error propagation. The Eq. (10) indicates that 'λ'_{SSIM}' is adaptively adjusted to be smaller than 'λ_{SSIM}' with the different packet loss rates to promote the error robustness of the video streaming. However, the work in [2] does not provide queue management for data traffic flow in the wireless network for congestion control.

3.3 Approach B: Queue Management Scheme Using CA-AQM [3]

To optimize the data traffic flow modeling, a cross layer optimization of video stream traffic at the router level was proposed in [3]. Here, coding was introduced at the network abstraction layer (NAL), where the queue based congestion control following Active Queue Management (AQM), and relative quality of Service (QoS) is mapped to schedule the rate of traffic flow. The cross layer approach is called as the CA-AQM process on a measured queue length, and derives the packet enqueue or dropping probability based on the receiving data traffic. At the Video coding layer (VCL), the video source is placed into slices and passed to the NAL for rate allocation. The method computes the drop probability dp(t) as;

$$dp(t) = \begin{cases} 0; & q_l(t) < min_{th} \\ 1; & q_l(t) > max_{th} \\ max_p \times \dfrac{q_l(t) - min_{th}}{max_{th} - min_{th}}; & \textit{otherwise} \end{cases} \tag{11}$$

where min_{th} is the minimum queue threshold, max_{th} is the maximum queue limit, and q_l is the current blockage factor. In the approach of cross layer modeling (CA-AQM), the drop probability is modified as; $dp(t) = 1 - \phi^{-p(t)}$, where p(t) is the price in period t and ϕ

constant value 1.001 defined as REM (random exponential marking). The price is updated from time to time according to the average queue length, the input rate, and the output rate of the queue. The CA-AQM approach controls the flow of traffic by accepting or dropping the video packets based on the probability index d(t) and the importance of the packet to drop. The price is incremented if the input rate exceeds the output rate, and is decremented otherwise. The suggested controlling algorithm of CA-AQM is outlined in Fig. 2.

Where, U(i) is the importance index of i packet in the queue. In the conventional queue mechanism (CA-AQM), it is seen that, the data traffic model is developed in observation to the current blockage level. This approach improves the network flow and intern network throughput, however, the effect of channel distortion on this traffic model is not evaluated. In other words, the allocation is not governed by the channel distortion level. Hence it can be inferred that, work of [3] does not cover the aspect of error resilient coding.

Therefore, it is imperative to have a multi attribute monitoring system, where the distortion level and the congestion factors are monitored simultaneously for the final allocation of data rate. To develop this concept, this paper proposes a novel data traffic approach at link level, called 'Dual Metric Traffic control' (DMTC), wherein two metrics, data congestion and distortion level are considered.

4 Dual Metric Traffic Control (DMTC)

To develop an optimal data traffic model with low blockage streaming, each link node is defined with a queue management scheme. Here, the data traffic flow is governed by the current congestion level per node and the rate control is applied with two quality metrics, blockage factor and data accuracy level.

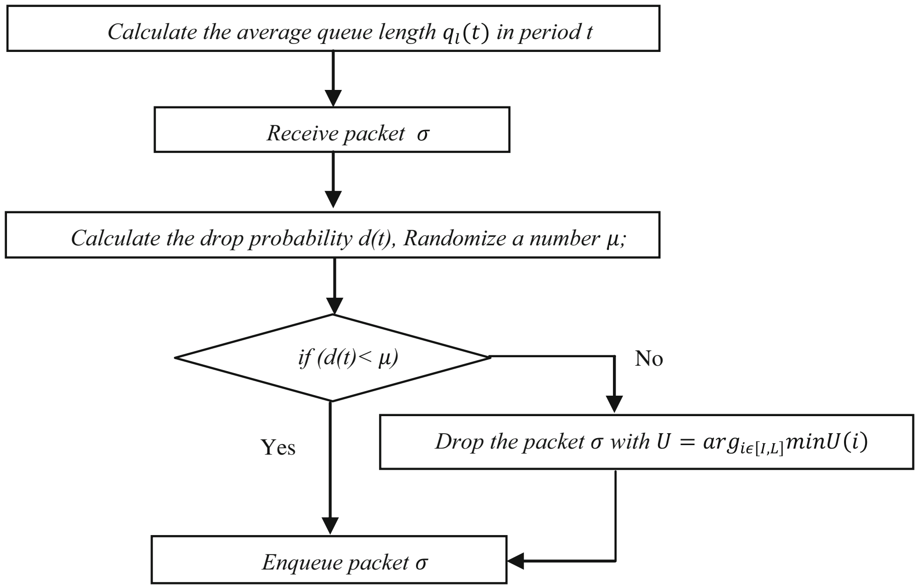


Fig. 2 Flow diagram for CA-AQM [3]

For a noise minimization in progressive streaming, iterative distortion estimate (IDE) was presented in [4]. In channel, noise effects are highly dynamic in nature. In this scenario, similarity measure factor for channel distortion estimation is non-effective. To overcome the problem of distortion diversity, the conventional SSIM based approach is modified to a cumulative distortion SSIM (CDSSIM) with iterative distortion measure and is given as,

$$CDSSIM = 1 - SSIM \tag{12}$$

To retain the loss control, the error estimate is defined on a per frame inter correlation basis. In the transmission of video frame, a group of block (GOB) coding is made. This approach defines the error values in each of the transmitting frame, where GOB are used to decode the error condition. The Errror (E) for a video sequence is defined by,

$$= P_0(D_1) + P_1(D_2) + P_2(D_3) \tag{13}$$

here P_i is the probability of the condition (D_i). D_i are defined as three possible conditions of GOB data, where D_1 reflect the accurate reception of current GOB, D_2 reflect the loss of current GOB but last GOB received, and D_3 reflect the loss of current and last GOB data. In the decoding process under these three cases, when the GOB are received correctly, the pixel data are reconstructed with minimal distortion DSSIM. While the case where current GOB is lost in channel, the last accurate GOB is taken as reference to rebuild the frame data. In case of both current and last GOB lost, cumulative distortion measured, CD-SSIM is then used and the pixels are estimated based on the minimization of CDSSIM factor. The Cumulative distortion SSIM reflects the distortion level, which are dynamically varied over a period of time. The integrated distortion estimate (IDE) in this case is defined as,

$$\mathrm{IDE} = (1 - \varepsilon)(D_1) + \varepsilon(1 - \varepsilon)(D_2) + \varepsilon^2(D_3) \tag{14}$$

where ε is the packet loss.

4.1 DMTC Approach

From the queue management technique, described in previous section, it is observed that, the congestion level is governed at two levels and the dropping probability is then defined as of '1' or '0' as presented in Eq. (11). It is also observed; traffic flow under $\min_{th}$ is considered as a non-congestive zone and above $\max_{th}$ is considered as a congestive zone. The region in between these two limits is taken as a random zone, where the packets have randomly been enqueued or dropped based on the probability of dropping $dp(t)$.

The rate of allocation (ROA) for the traffic flow in this case is presented as;

$$R_a(t) = \begin{cases} R_e(t) + \alpha t & \text{if} \quad Q_{cur} < \min_{th} \\ R_e(t) + (\alpha t - dp(t)) & \text{if} \quad \min_{th} < Q_{cur} < \max_{th} \\ R_e(t) - \dfrac{R_e(t)}{dp(t)} & \text{if} \quad Q_{cur} \geq \max_{th} \end{cases} \tag{15}$$

where $R_a(t)$ = data rate allocated, $R_e(t)$ = offered Data rate, αt = step of incremental data rate, Q_{cur} = current queue length, $\min_{th}$ = minimum queue limit, $\max_{th}$ = maximum queue limit, dp = dropping probability.

It can be seen from Eq. (15), the allocated data rate is varying with respect to the link blockage level. If the current queue length is below the minimum level, the rate is allocated with an increment of 'αt'. Where 'αt' is a constant increment factor defined by the network condition. If the current queue length is in between the minimum and maximum levels, representing partial blockage, the allocated data rate depends on the dropping probability (16). However, if the current queue length exceeds the maximum queue length, representing high congestive level, the allocated data rate is decreased as a fraction of drop probability. Here, the observation is, allocation of data rate is a function of blockage factor. To introduce the nonlinear distortion variations due to dynamic channel conditions, as second monitoring parameter, following possible cases arises:

4.2 Case 1: Under Invariant Channel Condition

In the transmission system, where the channel is invariant with time, a fixed scaling factor can be defined. According to the Eq. (15), the data rate allocated is dynamic in nature. This also effects the adjustment of the Lagrange multiplier in Eq. (8). The Lagrange multiplier in this case is defined by,

$$\lambda'_{SSIM-DMTC} = \frac{E}{R_{a(n,m,k)}(t)}$$

where E is defined as,

$$\begin{aligned} E &= \left(1 - \rho_{n,m} \cdot SSIM\left(b_{n,m,k}, b^{e_c}_{n,m,k}\right) - \left(1 - \rho_{n,m}\right) \cdot SSIM\left(b_{n,m,k}, b^{n_l}_{n,m,k}\right)\right) \\ &= \frac{\left(\rho_{n,m} \cdot SSIM\left(b_{n,m,k}, b^{e_c}_{n,m,k}\right)\right)}{R_{a(n,m,k)}(t)} + \frac{\left(\left(1 - \rho_{n,m}\right) \cdot SSIM\left(b_{n,m,k}, b^{n\_l}_{n,m,k}\right)\right)}{R_{a(n,m,k)}(t)} \end{aligned} \tag{16}$$

Approximately represented as;

$$\frac{\left(\rho_{n,m} \cdot SIM\left(b_{n,m,k}, b^{e\_c}_{n,m,k}\right)\right)}{R_{a(n,m,k)}(t)} = -\frac{\left(\left(1 - \rho_{n,m}\right) \cdot SSIM\left(b_{n,m,k}, b^{n_l}_{n,m,k}\right)\right)}{R_{a(n,m,k)}(t)} \approx \lambda_{SSIM} \tag{17}$$

From the above Eq. (17), it can be seen, the Lagrange Multiplier depends on the rate allocated for a particular nth slice in the mth frame. Thus, by adjusting the allocated data rate, the Lagrange multiplier is also adjusted, and provides an efficient error resilient and congestion free coding. From the Eq. (17), the SSIM-based Lagrange multiplier λ_{SSIM} in Eq. (5) can be modified as;

$$\lambda_{SSIM\_DMTC} = \begin{cases} \dfrac{-D_{SSIM}}{R_e(t) + \alpha t} & \text{if} \quad Q_{cur} < min_{th} \\ \dfrac{-D_{SSIM}}{R_e(t) + (\alpha t - dp(t))} & \text{if} \quad min_{th} < Q_{cur} < max_{th} \\ \dfrac{-D_{SSIM}}{R_e(t) - \dfrac{R_e(t)}{dp(t)}} & \text{if} \quad Q_{cur} \geq max_{th} \end{cases} \tag{18}$$

From the above Eq. (18), it is clear that the adjustment of the Lagrange multiplier in Lagrange optimization depends on the allocated data rate. For the measured distortion DSSIM and for the measured Lagrange multiplier λ_{SSIM_DMTC}, the Lagrange Optimization scheme (2), can be modified as

$$\min_{\{m\}}\{J\} = D_{SSIM} + \lambda_{SSIM\_DMTC} \cdot R_e \tag{19}$$

where 'DSSIM' denotes the SSIM-based distortion and 'λ_{SSIM_DMTC}' is the Lagrange multiplier. The Dual measuring metric of distortion monitoring and data rate allocation is made here based on the current queue length. The Dual metric monitoring hence guarantee a higher network throughput with better coding accuracy in video monitoring. The operational flow chart for the given approach is illustrated in Fig. 3.

4.3 Case 2: Under Variant Channel Condition

In the dynamic interference condition, the integrated distortion estimation is used for the optimization of Lagrange function, defined by the evaluation of network estimate. For the minimization of distortion (under variant noise condition) an optimization regression model, which minimizes the input–output based residual, is derived. The regression coefficient obtained through the sum of absolute value is then defined as [4],

$$\min_{W_0,W}\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(\mathrm{CDSSIM} - W_0 - W^T I_i\right)^2 + \lambda||W||\right\} \tag{20}$$

where 'n' represents the total number of blocks in GOB. The term w is the vector of regression coefficients, w0 is the intercept and λ is the regularization parameter. Substituting the Lagrange regularize parameter, the regression coefficient is then defined by,

$$\min_{W_0,\,W}\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(\mathrm{CDSSIM} - W_0 - W^T I_i\right)^2 + \lambda_{SSIM'}||W||\right\} \tag{21}$$

where λ'_{SSIM} is the regularizing parameter, using similarity index measure at the allocated transmission rate, and CDSSIM reflects the distortion measured over a period of observation. The proposed approach has the notion of dual metric observations for distortion minimization, wherein the similarity measure is used as a measuring parameter for rate allocation using λ'_{SSIM} and W_0, W to optimize CDSSIM. Hence, the dual metric optimization achieves the data rate allocation and distortion minimization under dynamic noise condition.

The rate allocation approach (ROA) under variant channel condition is dependent on the cumulative error function. It is defined by the optimization of regression parameter, where the minimization of cumulative distortion error due to Similarity index is carried out. The Lagrange regulator and the CDSSIM parameters are observed for rate allocation.

The stated rate allocation is then defined as,

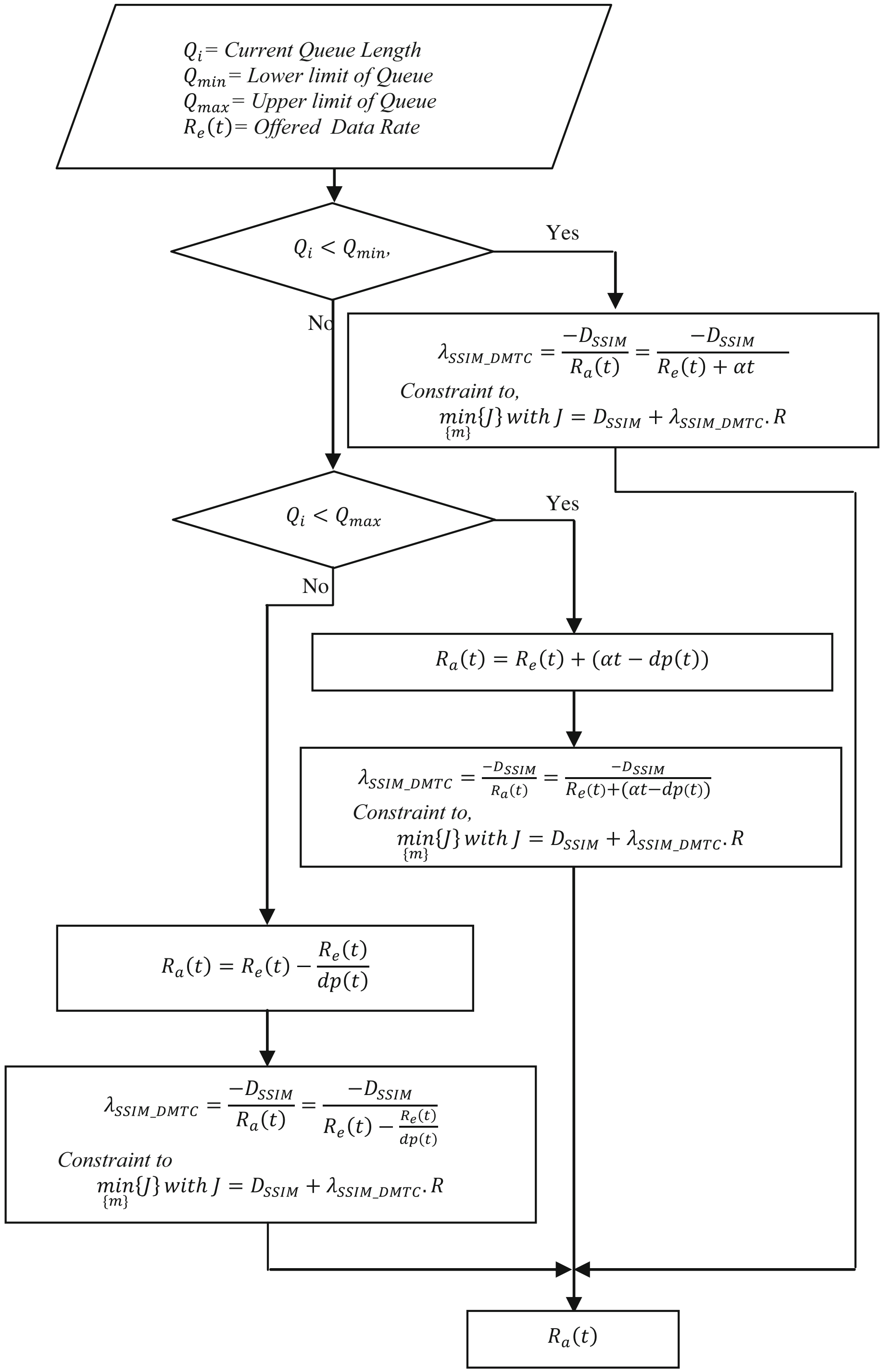


Fig. 3 Operation flow chart for the proposed SSIM based rate allocation approach

$$\lambda_{SSIM\_DMTC} = \left\{ \begin{array}{lll} \dfrac{-\,CDSSIM}{R_e(t) + \alpha t} & \text{if} & Q_{cur} < min_{th} \Rightarrow \min\limits_{W_0, W} \left\{ \dfrac{1}{2n} \sum\limits_{i=1}^{n} \left(CDSSIM - W_0 - W^T I_i\right)^2 + \lambda'_{SSIM} ||W|| \right\} \\ \dfrac{-\,CDSSIM}{R_e(t) - \alpha t} & \text{if} & Q_{cur} < min_{th} \not\Rightarrow \min\limits_{W_0, W} \left\{ \dfrac{1}{2n} \sum\limits_{i=1}^{n} \left(CDSSIM - W_0 - W^T I_i\right)^2 + \lambda'_{SSIM} ||W|| \right\} \\ \dfrac{-\,CDSSIM}{R_e(t) + (\alpha t - dp(t))} & \text{if} & min_{th} < Q_{cur} < max_{th} \Rightarrow \min\limits_{W_0, W} \left\{ \dfrac{1}{2n} \sum\limits_{i=1}^{n} \left(CDSSIM - W_0 - W^T I_i\right)^2 + \lambda'_{SSIM} ||W|| \right\} \\ \dfrac{-\,CDSSIM}{R_e(t) - (\alpha t - dp(t))} & \text{if} & min_{th} < Q_{cur} < max_{th} \not\Rightarrow \min\limits_{W_0, W} \left\{ \dfrac{1}{2n} \sum\limits_{i=1}^{n} \left(CDSSIM - W_0 - W^T I_i\right)^2 + \lambda'_{SSIM} ||W|| \right\} \\ \dfrac{-\,CDSSIM}{R_e(t) - \dfrac{R_e(t)}{dp(t)}} & \text{if} & Q_{cur} \geq max_{th} \Rightarrow \min\limits_{W_0, W} \left\{ \dfrac{1}{2n} \sum\limits_{i=1}^{n} \left(CDSSIM - W_0 - W^T I_i\right)^2 + \lambda'_{SSIM} ||W|| \right\} \\ 0, & \text{if} & Q_{cur} \geq max_{th} \not\Rightarrow \min\limits_{W_0, W} \left\{ \dfrac{1}{2n} \sum\limits_{i=1}^{n} \left(CDSSIM - W_0 - W^T I_i\right)^2 + \lambda'_{SSIM} ||W|| \right\} \end{array} \right\} \tag{22}$$

Here, the allocation problem is defined as a optimization of Lagrange function λ'_{SSIM} which is a function of allocation data rate w.r.t. similarity index measure. In the dynamic conditions the variant is measured as a factor of cumulative distortion measure (CDSSIM) which is also to optimize for rate allocation. In case the minimization cost function is satisfied, the rate allocation is subjected to increase by a factor of αt under the constraint of minimum threshold. In the same case if the regression model doesn't obtained to an optimization value, the allocation rate is decreased to achieve the convergence of minimum distortion. Under the intermediate region the data are dropped in a random manner based on the drop probability and the allocation is controlled, subjected to the minimization of regression error. The similar process is made with the maximum bound limit under two observing cases. Here, the data traffic is totally closed under the condition of convergence not meeting to the minimization criterion. The dual monitoring factor results in maximum accuracy and higher throughput under dynamic channel condition. Here the cumulative distortion results in minimization of distortion in channel variant condition.

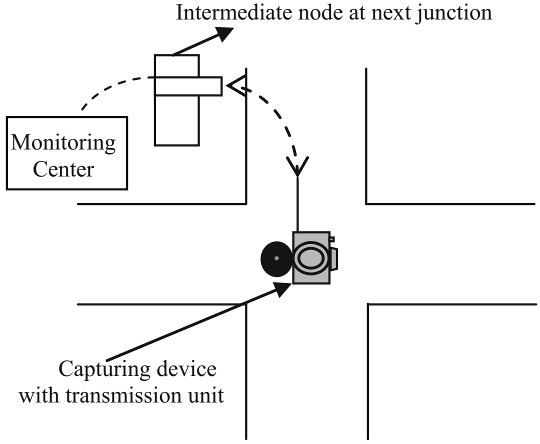
5 Experimental Results

To simulate the proposed approach a video compression algorithm at the video coding layer (VCL) is used, which is developed in [1]. The illustration of the communication model is shown in Fig. 4.

The constituting unit for the traffic surveillance approach is illustrated in Fig. 5.

At the video coding layer, the captured video is processed for compression. The motion elements are extracted using a recurrent block matching approach. The extracted motion vectors are compressed using an entropy encoder and stream out to Network abstraction layer (NAL). At each node, the NAL computes the current congestion level and computes the allocable rate of transmission in consideration with the error factor as briefed earlier. To evaluate the proposed approach, a subjective and objective analysis of the developed system is done with the approach of the SSIM based Rate allocation method. The performance of the proposed approach was evaluated with respect to SSIM, throughput, node overhead, end-to-end delay, and the allocated data rate. A network layout with a capturing

Fig. 4 Communication model for traffic surveillance



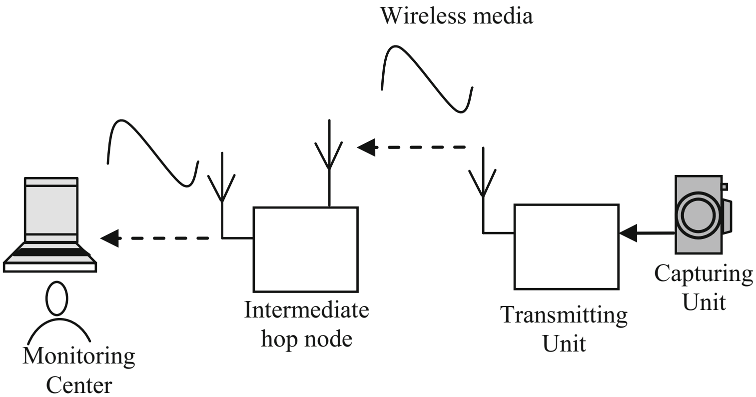


Fig. 5 Operational data flow for traffic surveillance

node, two intermediate hop nodes and a monitoring node is developed as illustrated in Fig. 6.

The network parameter used for the communication model is defined in Table 1.

The Fig. 7 illustrates the captured sequence from a traffic junction. The capturing unit was installed at the existing traffic light poles with a rotation of 360 degree orientation. The video is captured from a high resolution camera at a frame rate of 25 fps, with a 272×352 pixel resolution.

To process the video sample, the captured video sequence is extracted as frames. The frames are extracted at a skip of ten frames to ease the computation overhead. The extracted frame for processing is shown in Fig. 8.

Frames are recovered using conventional SSIM-RDO, but without data flow control (Fig. 9).

Frames are recovered using SSIM-RDO along with data flow control (Non-variant channel conditions) (Fig. 10).

Frames are recovered using SSIM-RDO along with data flow control (Variant channel conditions) (Fig. 11).

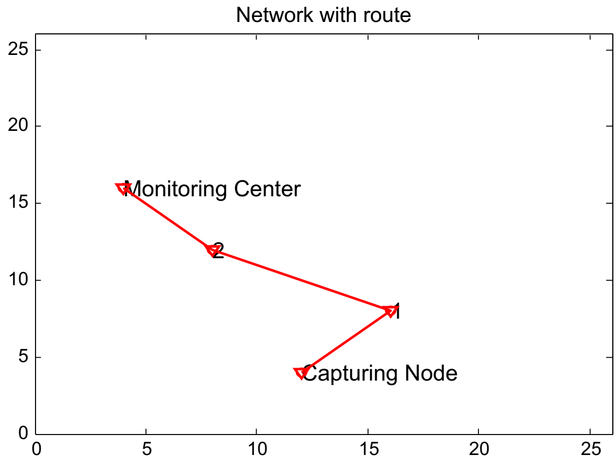


Fig. 6 Network model used for simulation

Table 1 Network parameter used

Network parameter	Values
Node placement	Static
Transmission range	40 units
Network area	25 m × 25 m
Number of nodes	4
Memory size	3 M
Qmin	0.15 × M
Qmax	0.75 × M
Initial blockage probability	0.1

Fig. 7 Captured surveillance video data

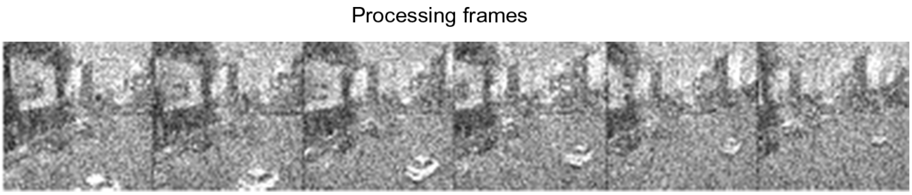


Fig. 8 Processing frames for the captured video sequence

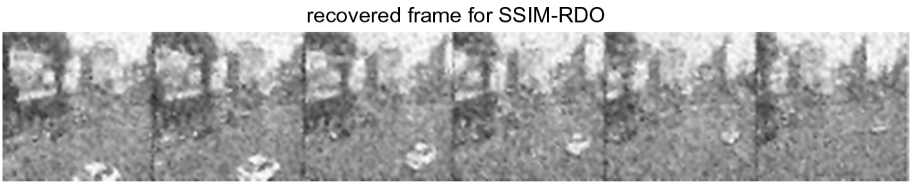


Fig. 9 Recovered frame at receiver using SSIM-RDO approach

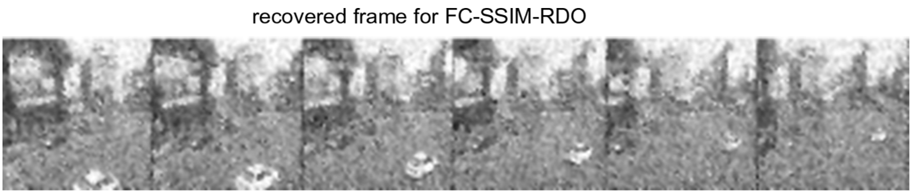


Fig. 10 Recovered frame using flow control approach

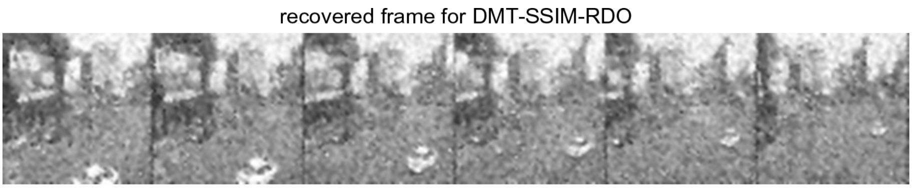


Fig. 11 Recovered frame using dual metric traffic control approach

Figures 12, 13, 14 and 15 show data transmission quality metrics i.e. route overhead, throughput, end to end delay and allocated data rate of SSIM-RDO, flow control (DMTC without channel noise) and DMTC approach. In this simulation, we considered ideal channel i.e. 0% noise. The graph shows that the performance of DMTC is better than existing approach of SSIM-RDO mechanism because of data flow control mechanism of DMTC.

5.1 Observation Under Variant Channel Conditions

Case 1 Variance = 0.1 (Figs. 16, 17, 18, 19).

Figures 20, 21, 22 and 23 show route overhead, throughput, end to end delay and allocated data rate of SSIM-RDO, flow control (DMTC without channel noise) and DMTC

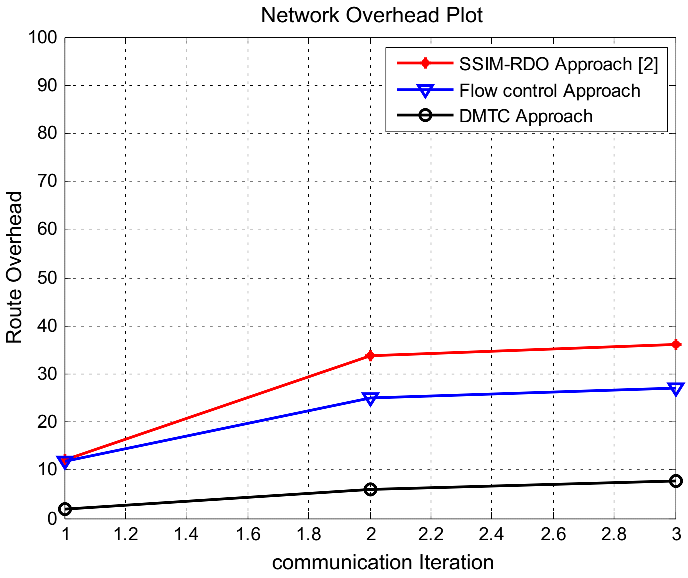


Fig. 12 Network overhead plot under non-variant channel condition

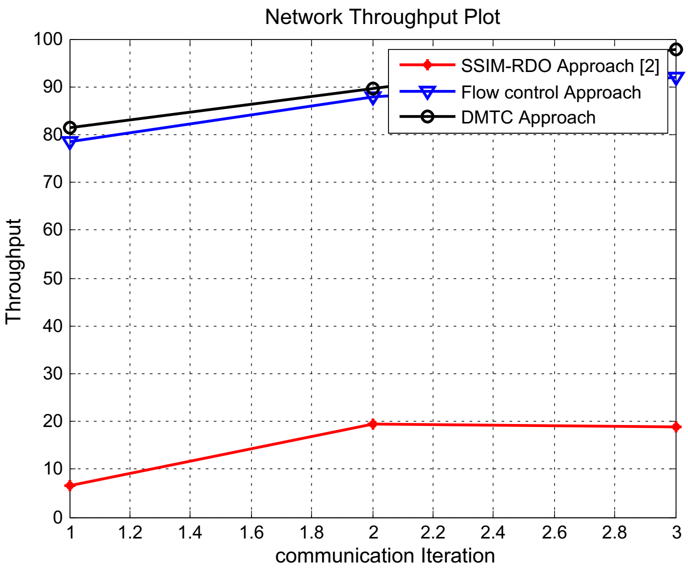


Fig. 13 Throughput plot for the developed approaches under non-variant channel condition

approach. In this simulation, we considered channel noise level at 10%. The graph shows that the performance of DMTC is better than existing approach of SSIM-RDO because of effective data flow control in DMTC approach. There is some improvement in DMTC from its counterpart i.e. flow control (without noise consideration).

Case 2 Varaince = 0.3 (Figs. 24, 25, 26, 27).

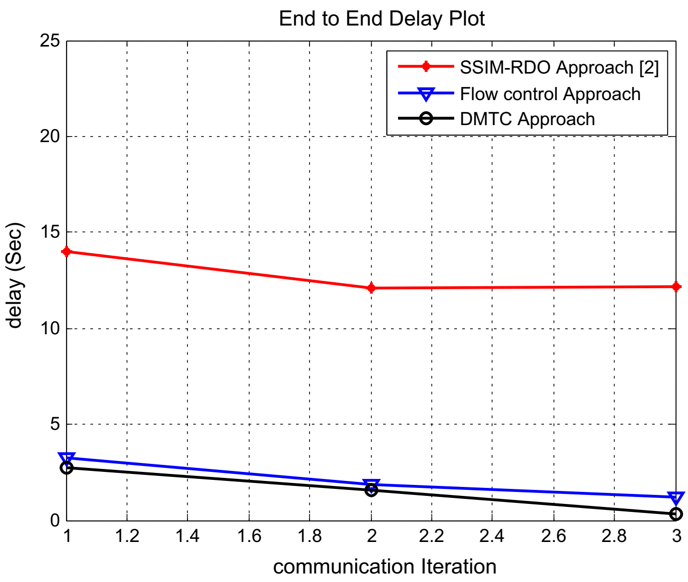


Fig. 14 End to end delay for developed approaches under non-variant condition

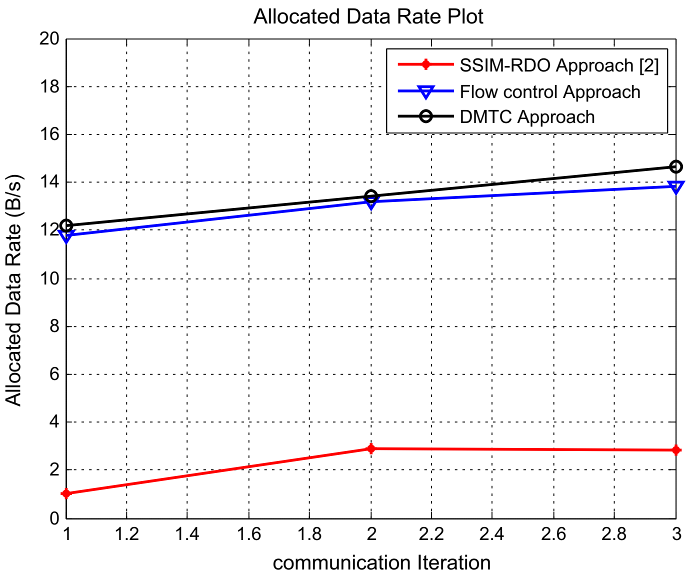


Fig. 15 Allocated data rate plot for developed approaches under non-variant condition

Figures 28, 29, 30 and 31 show route overhead, throughput, end to end delay and allocated data rate of SSIM-RDO, flow control (DMTC without channel noise) and DMTC approach. In this simulation, we considered channel noise level at 30%. The graph shows that the performance of DMTC is better than existing approach of SSIM-RDO because of effective data flow control in DMTC approach. Also, there is some improvement in DMTC from its counterpart i.e. flow control without noise consideration. Normally, increase in

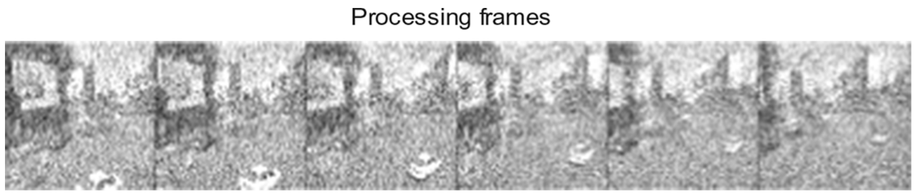


Fig. 16 Noised sample at channel variance = 0.1

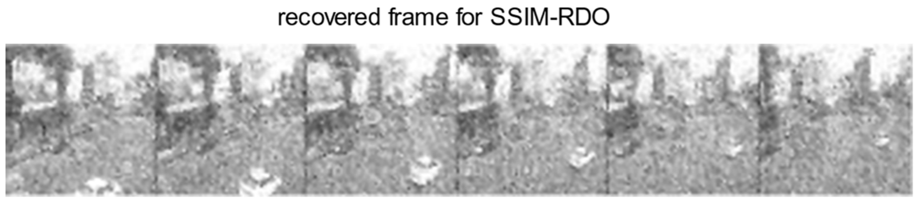


Fig. 17 Recovered sample at variance = 0.1 using SSIM approach

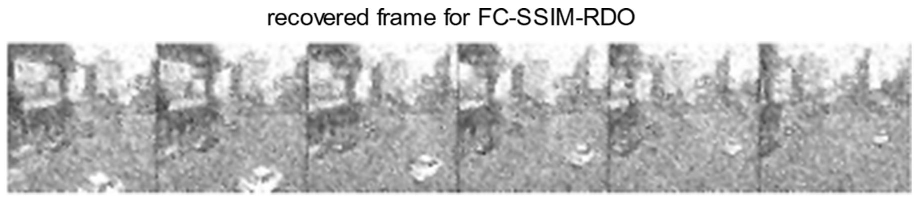


Fig. 18 Recovered sample at variance = 0.1 using flow control approach

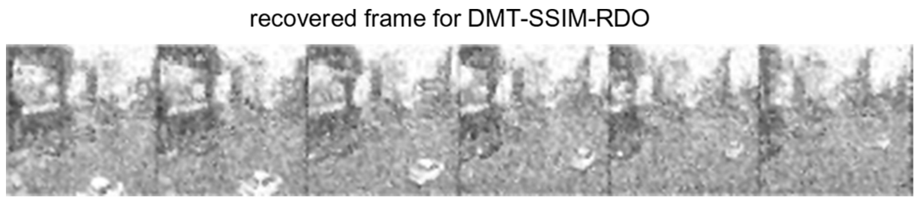


Fig. 19 Recovered sample at variance = 0.1 using DTMC approach

channel noise affects aforementioned quality metrics. But, it is evident from figures, DMTC approach does not let it to decrease, rather it is almost constant. Therefore, we term this as an improvement in the performance.

6 Conclusions and Future Scope

Novelty of proposed DMTC mechanism is an integration of the data traffic congestion metric with SSIM-RDO under variable and non variable channel conditions. The approach of error resilience with a high data traffic flow under channel variant condition is presented.

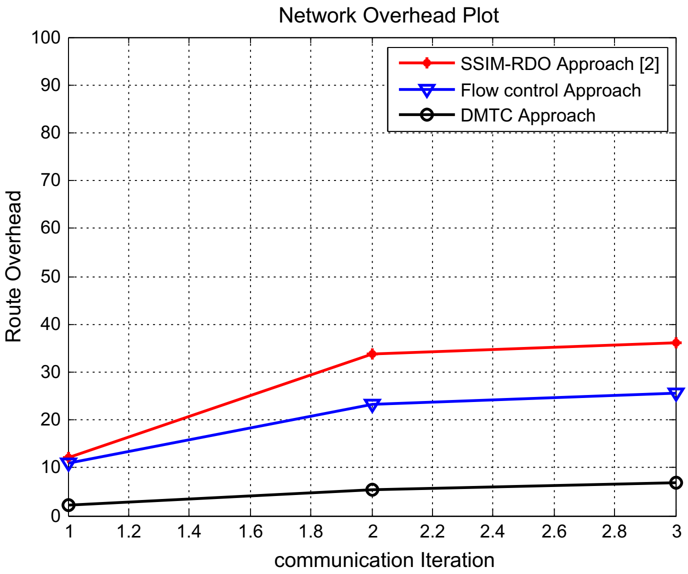


Fig. 20 Route overhead plot at variance = 0.1

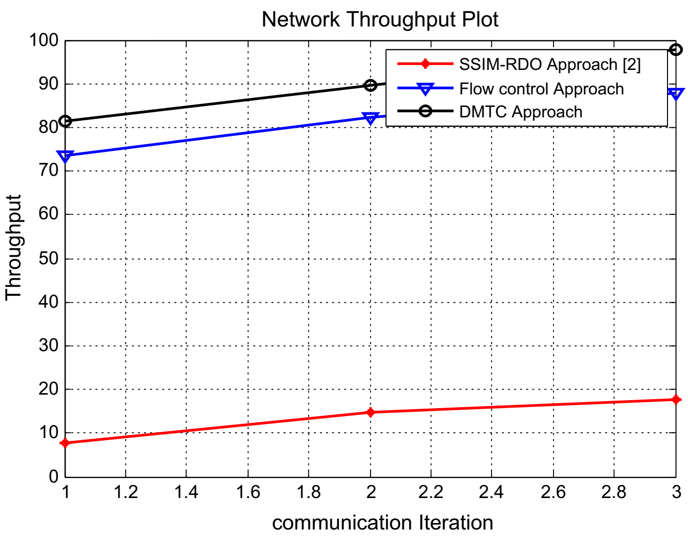


Fig. 21 Network throughput plot at variance = 0.1

Also, dynamic data flow control model with probabilistic route density is developed, to control the flow of the captured video data over a multi-hop wireless network model. In this approach, the video quality improvement is achieved with error resilience coding using the SSIM factor. The error resilience coding is then improved for high throughput using data flow controlling through rate allocation approach under variant channel condition.

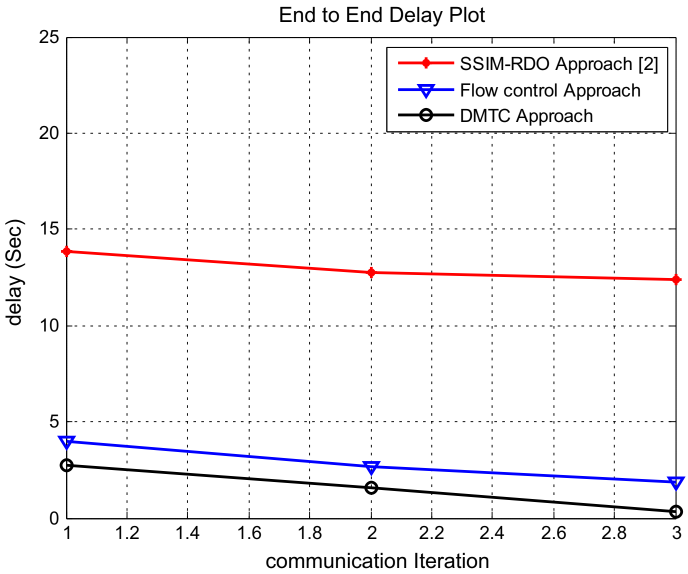


Fig. 22 End-to-end delay plot for channel variance = 0.1

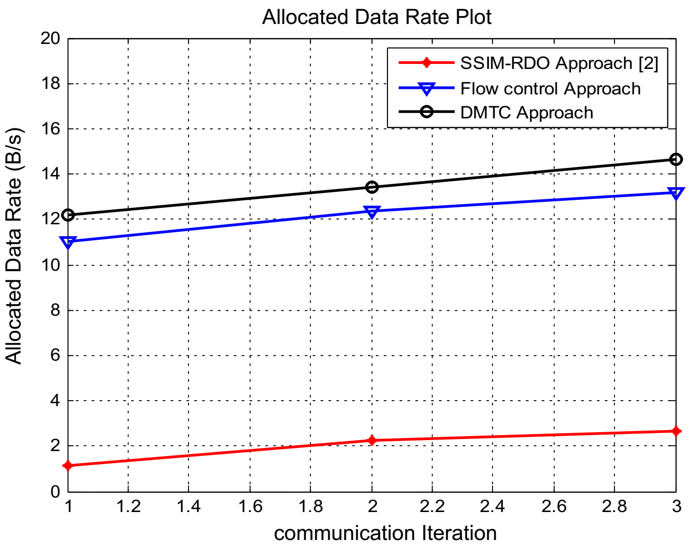


Fig. 23 Allocated data rate plot at variance = 0.1

From the experimental results, it is evident that an improvement in system throughput along with video quality is achieved.

Despite the effectiveness of proposed algorithm, there is a scope for further improvement. With improved variable block size segmentation or with improved motion vector prediction, work can be extended for H.265 codec in place of H.264, thereby increasing data compression and reducing communication overheads.

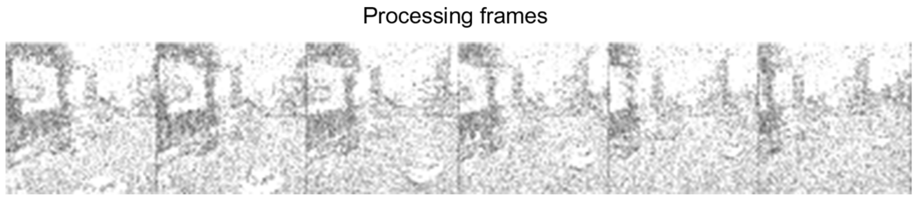


Fig. 24 Noised sample at channel variance = 0.3

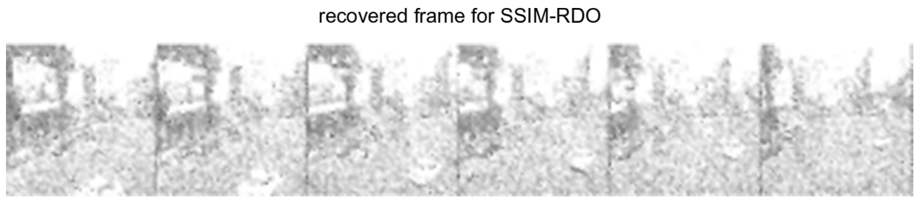


Fig. 25 Recovered sample at variance = 0.3 using SSIM approach

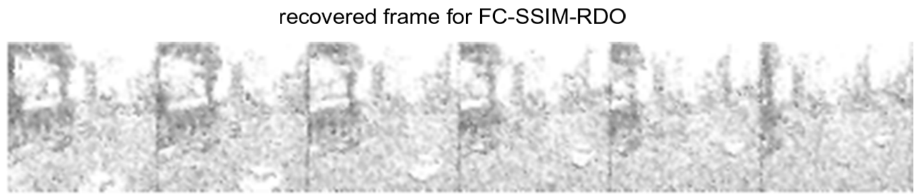


Fig. 26 Recovered sample at variance = 0.3 using flow control approach

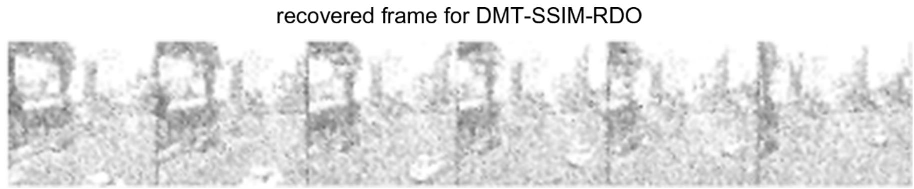


Fig. 27 Recovered sample at variance = 0.3 using DTMC approach

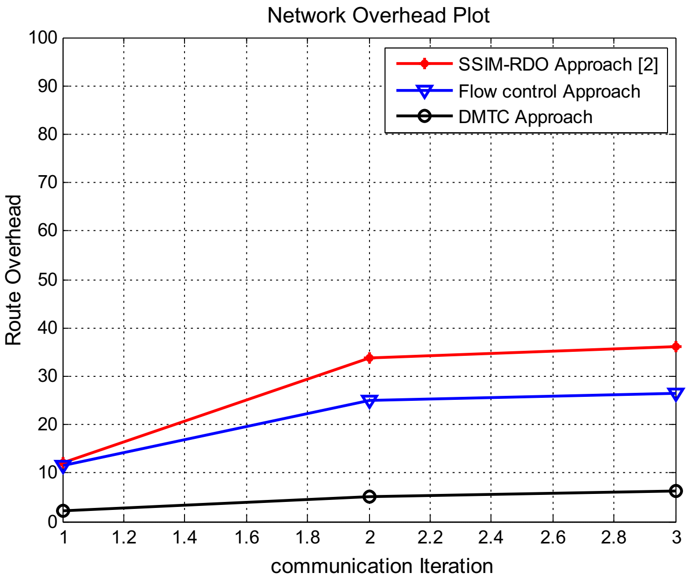


Fig. 28 Route overhead plot at channel variance = 0.3

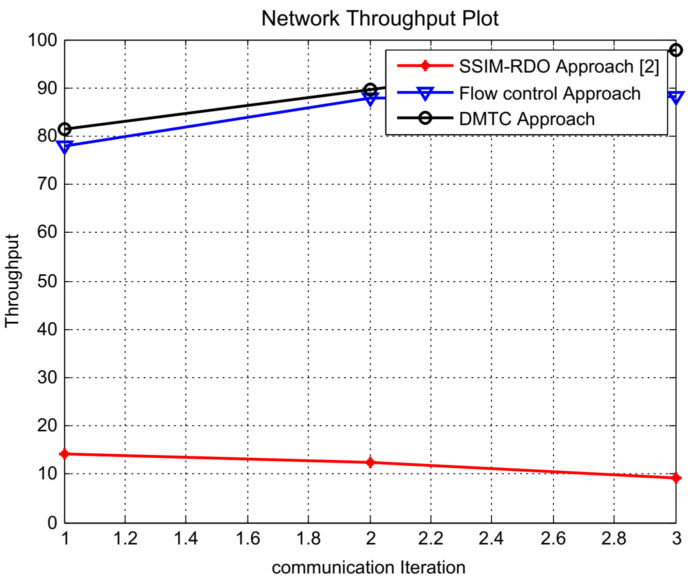


Fig. 29 Network throughput plot at channel variance = 0.3

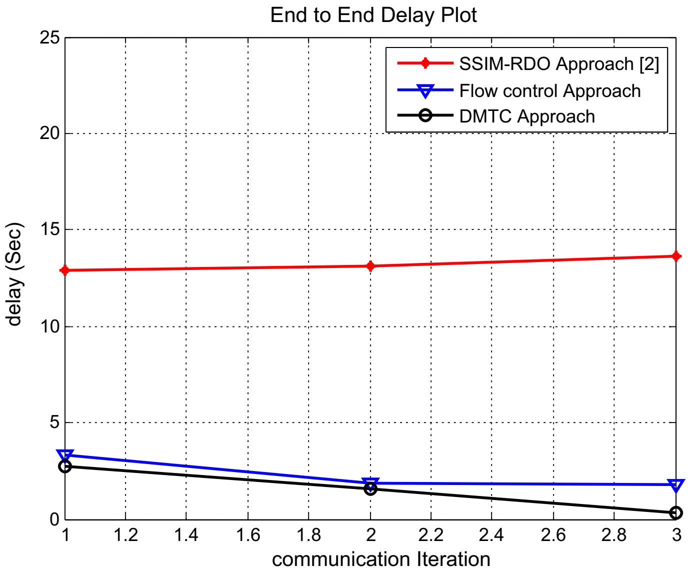


Fig. 30 End-to-end delay plot at channel variance = 0.3

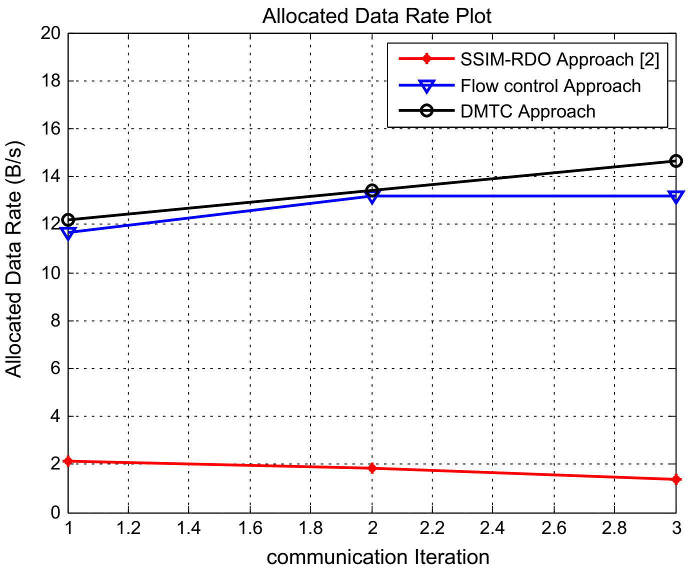


Fig. 31 Allocated data rate plot at variance = 0.3

References

1. Patil, S., Sanyal, R., & Prasad, R. (2015). Efficient video coding in region prediction in online video surveillance. In *The 2015 International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV)*, pp. 210–216.
2. Zhao, P., Liu, Y., Liu, J., Ci, S., & Yao, R. (2014). SSIM-based error-resilient rate-distortion optimization of H.264/AVC video coding for wireless streaming. *Signal Processing: Image Communication, 29,* 303–315.
3. Zhao, M. C., Gong, X. Y., Que, X. R., Wang, W. D., & Cheng, S. D. (2012). Context-aware adaptive active queue management mechanism for improving video transmission over IEEE 802.11E WLAN. *The Journal of China Universities of Posts and Telecommunications, 19*(Suppl. 2), 65–72.
4. Sankisa, A., Pandremmenou, K., Pahalawatta, P. V., Kondi, L. P., & Katsaggelos, A. K. (2016). SSIM-based distortion estimation for optimized video transmission over inherently noisy channels. *International Journal of Multimedia Data Engineering and Management, 7*(3), 34–52.
5. Ukhanova, A., Belyaev, E., Wang, L., & Forchhammer, S. (2012). Power consumption analysis of constant bit rate video transmission over 3G networks. *Computer Communications, 35,* 1695–1706.
6. Rezende, C., Mammeri, A., Boukerche, A., & Loureiro, A. A. (2014). A receiver-based video dissemination solution for vehicular networks with content transmissions decoupled from relay node selection. *Ad Hoc Networks, 17,* 1–17.
7. Nejad, A. E., & Romouzi, M. (2014). Investigation of video streaming application: a practical guidance. *Centre for Info Bio Technology, Indian Journal of Fundamental and Applied Life Sciences, 4*(S4), 399–407.
8. Qadri, N. N., Fleury, M., Altaf, M., & Ghanbari, M. (2010). Multi-source video streaming in a wireless vehicular ad hoc network. *The Institution of Engineering and Technology, 4,* 1300–1311.
9. Kazemian, H. B., & Ouazzane, K. (2013). Neuro-Fuzzy approach to video transmission over ZigBee. *Neurocomputing, 104,* 127–137.
10. Koutsia, A., Semertzidis, T., Dimitropoulos, K., Grammalidis, N., & Georgouleas, K. (2008). Intelligent Traffic Monitoring and Surveillance with Multiple Cameras. IEEE.
11. Hassan, M. M., & Hoong, P. K. (2013). Seamless handover integrated solution for video transmission over proxy mobile IPv6 in a micro mobility domain. *Journal of Network and Computer Applications, 36,* 66–76.
12. Lee, S., & Chung, K. (2008). Combining the rate adaptation and quality adaptation schemes for wireless video streaming. *Journal of Visual Communication and Image Representation, 19,* 508–519.
13. Parameswaran, V., Kannur, A., & Li, B. (2009). Adapting quantization offset in multiple description coding for error resilient video transmission. *Journal of Visual Communication and Image Representation, 20,* 491–503.
14. Lei, Z., & Georganas, N. D. (2005). Adaptive video transcoding and streaming over wireless channels. *The Journal of Systems and Software, 75,* 253–270.
15. Xing, M., & Cai, L. (2012). Adaptive video streaming with inter-vehicle relay for highway VANET scenario. In *IEEE ICC—Wireless Networks Symposium.*
16. Moorthy, A. K., & Bovik, A. C. (2009). A motion compensated approach to video quality assessment. In *IEEE Asilomar Conference on Signals, Systems and Computers,* pp. 872–875.
17. Seshadrinathan, K., & Bovik, A. C. (2007). A structural similarity metric for video based on motion Models. In *IEEE International Conference on Acoustics, Speech and Signal Processing* (Vol. 1, pp. 869–72), Honolulu, HI, USA.

Shivprasad P. Patil (M'66) received B.E. degree in Electronics Engineering from University of Pune, India, in 1989 and Master degree from Swami Ramanad Tirth Marathwada University, Nanded, in 2000. He is working as a Professor in the department of Information Technology in NBN Sinhgad School of Engineering, Pune. He is pursuing his Ph.D. from Aarhus University, Denmark. He has published 06 papers in international journals and conference proceedings in US and India. His research interests are in the areas of computer vision, multimedia data analysis and wireless multimedia communications.

Rajarshi Sanyal is telecommunications architect in the field of mobile communication. He commenced his career as a telecom engineer in Hutchison in the mobile switching domain in India and later on worked on R&D projects on SS7 and CTI applications. He was involved in R&D in a French Telecom startup in the role of a Domain Specialist-Telecom focusing on the development of new solutions related to CAMEL, SMSCs, Mobile Service Nodes and Mobile Payment solutions etc. He was associated with Reliance Communications as Engineering Manager on Mobile Core Networks and later on as Managing Consultant at IBM responsible for designing the Service Delivery Platform Solutions for Mobile Networks. He is presently working as a Core Network Architect (Voice and Mobile Data network engineering) at BICS (Belgacom International Carrier Services) at Brussels, Belgium. For the last 9 years at BICS, he had been spearheading tier one projects based on roaming hubs, GTP solutions, LTE instant roaming, Voice over LTE, WebRTC, M2M and IoT technologies, 5G, Mobile edge computing. He holds several EU, US and Indian patents in the field of Mobile and Computer communications. He has published 26 papers in international telecom journals and conference proceedings in US, Europe and Far East. He holds a Ph.D. from Aalborg University, Denmark. In his thesis he proposed a new generation multiple access technology based on intelligent physical layer to conceive a smart mobile network and realize device to device communication within the framework of next generation mobile network architecture.

Ramjee Prasad is a Professor of Future Technologies for Business Ecosystem Innovation (FT4B1) in the Department of Business Development and Technology, Aarhus University, Denmark. He is the Founder President of the CTIF Global Capsule (CGC). He is also the Founder Chairman of the Global ICT Standardisation Forum for India, established in 2009. GISFI has the purpose of increasing of the collaboration between European, Indian, Japanese, North-American and other worldwide standardization activities in the area of Information and Communication Technology (ICT) and related application areas. He has been honored by the University of Rome "Tor Vergata", Italy as a Distinguished Professor of the Department of Clinical Sciences and Translational Medicine on March 15, 2016. He is Honorary Professor of University of Cape Town, South Africa, and University of KwaZulu-Natal, South Africa. He has received Ridderkorset of Dannebrogordenen (Knight of the Dannebrog) in 2010 from the Danish Queen for the internationalization of top-class telecommunication research and education. He has received several international awards such as: IEEE Communications Society Wireless Communications Technical Committee Recognition Award in 2003 for making contribution in the field of "Personal, Wireless and Mobile Systems and Networks". Telenor's Research Award in

2005 for impressive merits both academic and organizational within the field of wireless and personal communication, 2014 IEEE AESS Outstanding Organizational Leadership Award for: “Organizational Leadership in developing and globalizing the CTIF (Center for TeleInFrastruktur) Research Network”, and so on. He has been Project Coordinator of several EC projects namely, MAGNET, MAGNET Beyond, eWALL and so on. He has published more than 30 books, 1000 plus journal and conference publications, more than IS patents, over 100 Ph.D. Graduates and larger number of Masters (over 250). Several of his students are today worldwide telecommunication leaders themselves.